

Requested Patent: JP2001290669A

Title:

SYSTEM, DEVICE AND METHOD FOR MANAGING FAIL OVER AND COMPUTER  
READABLE RECORDING MEDIUM ;

Abstracted Patent: JP2001290669 ;

Publication Date: 2001-10-19 ;

Inventor(s): HAYASE HISAO ;

Applicant(s): NS SOLUTIONS CORP ;

Application Number: JP20000108261 20000410 ;

Priority Number(s): ;

IPC Classification: G06F11/20; G06F13/00; G06F15/177 ;

Equivalents: ;

**ABSTRACT:**

**PROBLEM TO BE SOLVED:** To succeed a service between servers in a short time, without hindering existent services. **SOLUTION:** In the fail over managing system provided with a server A, a server B and a service disk D inside a RAID 1 physically connected to these servers A and B, the server B is provided with a script file 55 previously storing link information concerning stringing between a physical address showing a physical connection relation with the service disk D and a logical address, showing a logical connection relation with the service disk D and a link processing part 53 for spreading a link with the service disk D on the basis of that link information. When the server B is to succeed a service (a) of the server A, by setting the logical address according to the previously stored link information, it is not necessary to perform heavy processing, such as scanning a device.

(19)日本国特許庁 (JP)

## (12) 公開特許公報 (A)

(11)特許出願公開番号

特開2001-290669

(P2001-290669A)

(43)公開日 平成13年10月19日(2001.10.19)

(51)Int.Cl. <sup>7</sup>	識別記号	FI	テームト <sup>*</sup> (参考)
G 0 6 F 11/20	3 1 0	G 0 6 F 11/20	3 1 0 E 5 B 0 3 4
13/00	3 0 1	13/00	3 0 1 K 5 B 0 4 5
	3 5 1		3 5 1 M 5 B 0 8 3
15/177	6 7 8	15/177	6 7 8 B 5 B 0 8 9

審査請求 未請求 請求項の数11 OL (全 10 頁)

(21)出願番号 特願2000-108261(P2000-108261)

(22)出願日 平成12年4月10日(2000.4.10)

(71)出願人 000191076

新日鉄ソリューションズ株式会社

東京都中央区新川2丁目20番15号

(72)発明者 早瀬 久雄

東京都千代田区大手町2-6-3 新日本

製鐵株式会社内

(74)代理人 100090273

弁理士 國分 孝悦

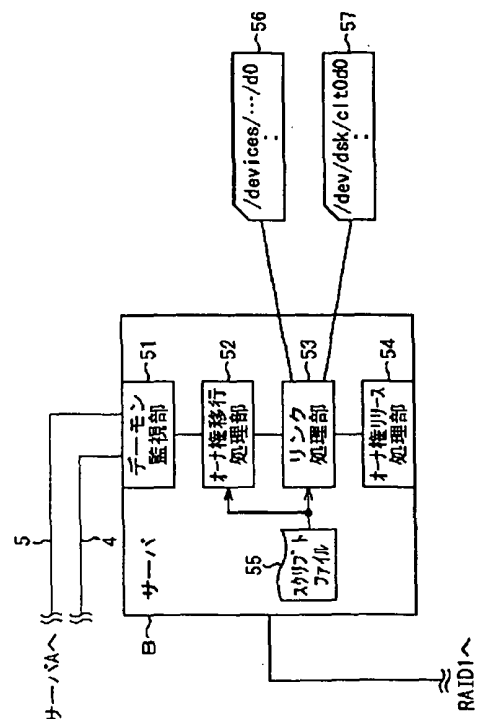
最終頁に続く

(54)【発明の名称】 フェールオーバー管理システム、フェールオーバー管理装置、方法、及びコンピュータ読み取り可能な記録媒体

## (57)【要約】

【課題】 短時間で、既存のサービスに支障を与えることなくサーバ間でサービスの引き継ぎが行えるようにする。

【解決手段】 サーバAと、サーバBと、これらサーバA、Bに物理的に接続されたRAID1内のサービスディスクDとを備えたフェールオーバー管理システムにおいて、サーバBに、サービスディスクDとの物理的接続関係を示した物理アドレスとサービスディスクDとの論理的接続関係を示した論理アドレスとのひもづけに関するリンク情報をあらかじめ格納したスクリプトファイル55と、そのリンク情報に基づいてサービスディスクDとのリンクをはるリンク処理部53とを設け、サーバBがサーバAのサービスaを引き継ぐ際に、事前に格納されているリンク情報に従って論理アドレスを設定することにより、デバイスをスキャンするような重い処理を行わなくても済むようになる。



## 【特許請求の範囲】

【請求項1】 少なくとも一のサーバと、他のサーバと、上記一のサーバ及び他のサーバに物理的に接続されたデバイスとを備え、上記デバイスのオーナー権を上記一のサーバから上記他のサーバに移行させて、上記一のサーバが提供するサービスを上記他のサーバに引き継がせる構成にしたフェールオーバー管理システムであって、上記他のサーバは、  
上記デバイスとの物理的接続関係を示した物理アドレスと上記デバイスとの論理的接続関係を示した論理アドレスとのひもづけに関するリンク情報をあらかじめ格納するリンク情報記憶手段と、  
特定の状態となったときに、上記リンク情報記憶手段に格納されている上記リンク情報に基づいて上記デバイスとのリンクをはるリンク処理手段とを備えたことを特徴とするフェールオーバー管理システム。

【請求項2】 上記一のサーバが上記デバイスのオーナー権を有して一のサービスを提供しているとき、上記他のサーバは待機中であることを特徴とする請求項1に記載のフェールオーバー管理システム。

【請求項3】 上記一のサーバが上記デバイスのオーナー権を有して一のサービスを提供しているとき、上記他のサーバは他のデバイスのオーナー権を有して他のサービスを提供していることを特徴とする請求項1に記載のフェールオーバー管理システム。

【請求項4】 上記他のサーバは、監視ネットワークを介して上記一のサーバを監視する監視手段を備え、上記リンク処理手段は、上記監視手段により上記一のサーバの障害発生が検出されたときに起動されることを特徴とする請求項1～3のいずれか1項に記載のフェールオーバー管理システム。

【請求項5】 上記他のサーバは、上記デバイスのオーナー権を自己に移行させる処理を行うオーナー権移行処理手段を備えたことを特徴とする請求項1～4のいずれか1項に記載のフェールオーバー管理システム。

【請求項6】 上記リンク情報記憶手段は、上記リンク情報をテーブル化して格納することを特徴とする請求項1～5のいずれか1項に記載のフェールオーバー管理システム。

【請求項7】 上記他のサーバは、上記デバイスのオーナー権をリリースする処理を行うオーナー権リリース処理手段を備えたことを特徴とする請求項1～6のいずれか1項に記載のフェールオーバー管理システム。

【請求項8】 少なくとも一のサーバと、他のサーバと、上記一のサーバ及び他のサーバに物理的に接続されたデバイスとを備えたフェールオーバー管理システムにおいて、上記デバイスのオーナー権を上記一のサーバから上記他のサーバに移行させて、上記一のサーバが提供するサービスを上記他のサーバに引き継がせるようにしたフェールオーバー管理方法であって、

特定の状態となったときに、上記他のサーバにあらかじめ記憶されている上記他のサーバと上記デバイスとの物理的接続関係を示した物理アドレスと上記他のサーバと上記デバイスとの論理的接続関係を示した論理アドレスとのひもづけに関するリンク情報に基づいて、上記他のサーバと上記デバイスとのリンクをはる手順を有することを特徴とするフェールオーバー管理方法。

【請求項9】 物理的に接続されたデバイスとの物理的接続関係を示した物理アドレスと上記デバイスとの論理的接続関係を示した論理アドレスとのひもづけに関するリンク情報をあらかじめ格納するリンク情報記憶手段と、

特定の状態となったときに、上記リンク情報記憶手段に格納されている上記リンク情報に基づいて上記デバイスとのリンクをはるリンク処理手段とを備えたことを特徴とするフェールオーバー管理装置。

【請求項10】 請求項1～7及び9のいずれか1項に記載の各手段としてコンピュータを機能させるためのプログラムを記録したことを特徴とするコンピュータ読み取り可能な記録媒体。

【請求項11】 請求項8に記載の各手順をコンピュータに実行させるためのプログラムを記録したことを特徴とするコンピュータ読み取り可能な記録媒体。

## 【発明の詳細な説明】

## 【0001】

【発明の属する技術分野】本発明は、一のサーバに障害が発生したときに、その一のサーバが提供していたサービスを他のサーバに引き継がせるようにしたフェールオーバー管理システム、フェールオーバー管理装置、フェールオーバー管理方法、さらにはこれらを実現するためのプログラムを記録したコンピュータ読み取り可能な記録媒体に関する。

## 【0002】

【従来の技術】サーバとクライアントとがネットワークを介して接続されたネットワークシステムにおいて、サーバの予期せぬシステムダウンは、業務停止を招き、業務の遂行効率を低下させてしまう。かかるサーバのシステムダウンによる業務停止を回避するため、複数台のサーバを接続したフェールオーバー管理システムが用いられている。このフェールオーバー管理システムでは、一のサーバに障害が発生したときに、その一のサーバが提供していたサービスを他のサーバに引き継がせるようにしている。

【0003】この種のフェールオーバー管理システムとしては、例えば、非対称型と呼ばれるタイプや対称型と呼ばれるタイプのものがある。図6に示すタイプは非対称型といわれるシステムであり、図6(a)に示すように、通常の状態では、サーバAがサービスディスクDを用いてサービスaを提供し、サーバBが待機中となっている。そして、サーバAに障害が発生してダウンした場

合には、図6(b)に示すように、待機中のサーバBがサービスディスクDのオーナー権を自動的に取得して、それまでサーバAが提供していたサービスaを引き継ぐようにしている。

【0004】また、図7に示すタイプは対称型といわれるシステムであり、図7(a)に示すように、通常の状態では、サーバAがサービスディスクDaを用いてサービスaを提供し、サーバBが別のサービスディスクDbを用いて別のサービスbを提供している。そして、サーバAに障害が発生してダウンした場合には、図7(b)に示すように、サーバBがサービスディスクDaのオーナー権を自動的に取得して、それまでサーバAが提供していたサービスaを引き継ぐようにしている。したがって、サーバBは、既存のサービスbとともに、引き継いだサービスaも提供することになる。同様に、サーバBに障害が発生してダウンした場合には、サーバAがサービスディスクDbのオーナー権を自動的に取得して、それまでサーバBが提供していたサービスbを引き継ぐようにしている。したがって、サーバAは、既存のサービスaとともに、引き継いだサービスbも提供することになる。

【0005】

【発明が解決しようとする課題】上記のように構成にしたフェールオーバー管理システムでは、サービスの引き継ぎを行うとき、単にディスクのオーナー権を移行するだけでは足りず、移行後に使用される各種デバイスを認識するためのデバイスファイルの再構築しなければならない。例えば、UNIX（登録商標）オペレーティングシステムの1つであるSolaris（米Sun Microsystems社）では、まず、`'drvconfig'`といったコマンドを使用して、接続されているSCSIのデバイスをすべてスキャンする。そして、`'disks'`といったコマンドを使用して、スキャンされたSCSIのデバイスの中からディスクを探してリンクをはることにしている。

【0006】しかしながら、上記デバイスファイルの再構築処理、特にSCSIのデバイスのスキャンは非常に重い処理で、その処理に非常に長い時間がかかってしまうといった問題があった。しかも、図7に示した対称型のシステムにおいては、一方のサーバでサービスを実行している途中で、他方のサーバのダウンによって非常に重いスキャン処理が行われ、サービスを引き継ぐ側のサーバに多大な負荷がかかるため、そのサーバが提供している既存のサービスにまで支障をきたす可能性があるといった問題もあった。

【0007】本発明は、このような問題を解決するために成されたものであり、短時間で、しかも、既存のサービスに支障を与えることなくサーバ間でサービスの引き継ぎを行えるようにすることを目的とする。

【0008】

【課題を解決するための手段】上記の課題を解決すべ

く、本発明のフェールオーバー管理システムは、少なくとも一のサーバと、他のサーバと、上記一のサーバ及び他のサーバに物理的に接続されたデバイスとを備え、上記デバイスのオーナー権を上記一のサーバから上記他のサーバに移行させて、上記一のサーバが提供するサービスを上記他のサーバに引き継がせる構成にしたフェールオーバー管理システムであって、上記他のサーバは、上記デバイスとの物理的接続関係を示した物理アドレスと上記デバイスとの論理的接続関係を示した論理アドレスとのひもづけに関するリンク情報をあらかじめ格納するリンク情報記憶手段と、特定の状態となったときに、上記リンク情報記憶手段に格納されている上記リンク情報に基づいて上記デバイスとのリンクをはるリンク処理手段とを備えた点に特徴を有する。

【0009】さらに、本発明のフェールオーバー管理システムは、上記一のサーバが上記デバイスのオーナー権を有して一のサービスを提供しているとき、上記他のサーバは待機中である点に特徴を有する。

【0010】さらに、本発明のフェールオーバー管理システムは、上記一のサーバが上記デバイスのオーナー権を有して一のサービスを提供しているとき、上記他のサーバは他のデバイスのオーナー権を有して他のサービスを提供している点に特徴を有する。

【0011】さらに、本発明のフェールオーバー管理システムは、上記他のサーバは、監視ネットワークを介して上記一のサーバを監視する監視手段を備え、上記リンク処理手段は、上記監視手段により上記一のサーバの障害発生が検出されたときに起動される点に特徴を有する。

【0012】さらに、本発明のフェールオーバー管理システムは、上記他のサーバは、上記デバイスのオーナー権を自己に移行させる処理を行うオーナー権移行処理手段を備えた点に特徴を有する。

【0013】さらに、本発明のフェールオーバー管理システムは、上記リンク情報記憶手段は、上記リンク情報をテーブル化して格納する点に特徴を有する。

【0014】さらに、本発明のフェールオーバー管理システムは、上記他のサーバは、上記デバイスのオーナー権をリリースする処理を行うオーナー権リリース処理手段を備えた点に特徴を有する。

【0015】また、本発明のフェールオーバー管理方法は、少なくとも一のサーバと、他のサーバと、上記一のサーバ及び他のサーバに物理的に接続されたデバイスとを備えたフェールオーバー管理システムにおいて、上記デバイスのオーナー権を上記一のサーバから上記他のサーバに移行させて、上記一のサーバが提供するサービスを上記他のサーバに引き継がせるようにしたフェールオーバー管理方法であって、特定の状態となったときに、上記他のサーバにあらかじめ記憶されている上記他のサーバと上記デバイスとの物理的接続関係を示した物理アドレスと上記他のサーバと上記デバイスとの論理的接続関係を

示した論理アドレスとのひもづけに関するリンク情報に基づいて、上記他のサーバと上記デバイスとのリンクをはる手順を有する点に特徴を有する。

【0016】また、本発明のフェールオーバー管理装置は、物理的に接続されたデバイスとの物理的接続関係を示した物理アドレスと上記デバイスとの論理的接続関係を示した論理アドレスとのひもづけに関するリンク情報をあらかじめ格納するリンク情報記憶手段と、特定の状態となったときに、上記リンク情報記憶手段に格納されている上記リンク情報に基づいて上記デバイスとのリンクをはるリンク処理手段とを備えた点に特徴を有する。

【0017】また、本発明のコンピュータ読み取り可能な記録媒体は、請求項1〜7及び9のいずれか1項に記載の各手段としてコンピュータを機能させるためのプログラムを記録した点に特徴を有する。

【0018】また、本発明のコンピュータ読み取り可能な記録媒体は、請求項8に記載の各手順をコンピュータに実行させるためのプログラムを記録した点に特徴を有する。

【0019】以上述べた本発明によれば、例えば一のサーバに障害が発生したとき、それまでそのサーバがデバイスのオーナー権を有して提供していたサービスを、他のサーバに引き継がせることができるので、サーバダウンに伴う業務停止を免れることができる。その際、サービスを引き継ぐ他のサーバでは、その引き継ぎ時に論理的に接続されるべきディスクの物理アドレスと論理アドレスとのひもづけがあらかじめなされているので、そのひもづけに関するリンク情報に基づいて、簡単に論理アドレスを設定することが可能となり、デバイスのスキャン等の重い処理を行わなくても済む。

【0020】

【発明の実施の形態】以下、図面に基づいて本発明の一実施形態を説明する。図1には、本実施形態による非対称型のフェールオーバー管理システムの概略構成を示す。このシステムでは、OS（オペレーティングシステム）として、UNIXオペレーティングシステムの1つであるSolaris（米Sun Microsystems社）を使用しているものとする。

【0021】1はRAID（redundant arrays of inexpensive disk）で、後述する2つのサーバA、Bに対して互いに独立した2つのRAIDコントローラ1a、1bを有する。RAIDコントローラをサーバA、Bに対して一つとすることも可能であるが、本実施形態のようにRAIDコントローラ1a、1bを互いに独立させて設けたのは、一方が故障しても、それに関係なく他方で対処でき、対故障性が高いためである。

【0022】上記RAID1には、サービスaを提供するためのサービスディスクDが内蔵されている。なお、図1では、サービスディスクDが1つのものとして表されているが、その個数や仕様はサービスaの内容等によ

って適宜変更されるものである。

【0023】上記サーバAは、RAID1のRAIDコントローラ1aに接続されている。通常は、このサーバAにサービスディスクDのオーナー権があり、サーバAがネットワーク2に対してサービスaを提供している。すなわち、サーバAの物理アドレスをc1、RAIDコントローラ1aのSCSIターゲットIDをt0、サービスディスクDの物理アドレスをd0と表せば、c1t0d0といった論理アドレスがサーバAに設定され、サービスディスクDを用いたサービスaがサーバAによって提供される。なお、このサーバAが、本発明でいう一のサーバを構成するものである。

【0024】また、上記サーバBは、RAID1のRAIDコントローラ1bに接続されている。通常は、このサーバBからはサービスディスクDが見えないように設定され、サーバBは待機中となっている。すなわち、サーバBの物理アドレスをc1、RAIDコントローラ1bのSCSIターゲットIDをt0と表せば、c1t0d0といった論理アドレスはサーバBに設定されず、サーバBからサービスディスクDにはアクセスできないようになっている。なお、このサーバBが、本発明でいう他のサーバを構成するものである。

【0025】3はRAID1内に設けた固定ディスクで、RAIDコントローラ1bを介して待機中のサーバBに接続されている。この固定ディスク3は、サービスディスクDのオーナー権移行処理を行うためサーバBにおいてコマンド“trespass”が発行されたとき、そのコマンドを受け取るのに必要となるものである。すなわち、このコマンド“trespass”は、ディスクに対して発行されるものであるが、通常の状態ではサーバBからサービスディスクDは見えないようになっているため、オーナー権を移行する際に仮にコマンド“trespass”を投げるダミーの固定ディスク3が必要となるのである。固定ディスク3の物理アドレスをd1で表すと、サーバBが待機状態にあるとき、論理アドレスはc1t0d1に設定されている。なお、図1では、固定ディスク3が1つのものとして表されているが、仮にこれが故障したときに冗長性を持たせるために2つ以上設けておくのが望ましい。

【0026】4、5は上記サーバAとサーバBとを接続する監視ネットワークで、これら監視ネットワーク4、5を介してサーバA、B間でデーモンに従った監視を行っている。なお、デーモンとは、システムに常駐して、アプリケーションやシステムが特定の状態になったときに自動的に各種のサービスを提供するプログラムのことである。具体的には、監視ネットワーク4では、ICMP（Internet Control Message Protocol）を用いて相手先に対して返答を要求するPING（Packet Internet Groper）を使用した通信を行い、監視ネットワーク5では、SOCKETを使用した通信を行っている。そし

て、待機中のサーバBは、設定時間内（例えば600sec）内に、監視ネットワーク4、5の両方で通信が途絶えたことを認定したとき、サービスaを提供しているサーバAに障害が発生したものと判別する。

【0027】このように監視ネットワーク4、5を二重化したのは、監視の信頼性を高めるためである。例えば、監視ネットワークを1系統しか設けなかった場合、この監視ネットワークで通信が途絶えても、サーバに障害が発生したのか、それとも監視ネットワーク自体に障害が発生したのかを判別することができない。それに対して、監視ネットワーク4、5のように監視デーモンを二重化しておけば、その両方で通信が途絶えた場合、両方の監視ネットワーク4、5自体で同時に障害が発生する（例えばハブ4a、5aが同時に故障する）ようなことはまず有りえないので、サーバAに障害が発生したものと判断することができる。

【0028】図2には、上記サーバBの機能構成を表すブロック図を示す。サーバBは、デーモン監視部51と、オーナ権移行処理部52と、リンク処理部53と、オーナ権リリース処理部54と、スクリプトファイル55の記憶部とを有する。

【0029】デーモン監視部51は、上述した監視ネットワーク4、5を介してサーバAの故障発生を監視するものである。オーナ権移行処理部52は、デーモン監視部51での監視結果に基づいて、上記サービスディスクDのオーナ権を自己に移行させる指令として、前述の“trespass”といったコマンドをスクリプトファイル55の記述に従って発行し、サービスディスクDのオーナ権をサーバBに移行させる処理を行うものである。

【0030】上記スクリプトファイル55には、上記オーナ権移行処理部52及びリンク処理部53で行われる処理を含む一連の処理の流れがスクリプト形式で記述されるとともに、上記サービスディスクDとサーバBとの物理的接続関係を示した物理アドレス56と上記サービスディスクDとサーバBとの論理的接続関係を示した論理アドレス57とのひもづけに関するリンク情報が記述されている。このリンク情報は、単純にリストの形で記述してもよいし、数が多くなるような場合はテーブルの形で記述してもよい。

【0031】上記物理アドレス56は、サーバBとサービスディスクDとのリンクに必要とされるサービスディスクDの物理的な位置を示すもので、システム導入時にわかっている。例えば、図2に示すものは、サービスディスクDの物理的な位置d0が定義されている。また、論理アドレス57は、サーバBとサービスディスクDとの論理的接続関係を示すもので、図2に示すように、サーバBとRAIDコントローラ1bとサービスディスクDとの関連c1t0d0が定義されている。

【0032】上述したように、サーバBが待機中のときは、サーバBは固定ディスク3と論理的に接続され、論

理アドレスはc1t0d1となっている。しかし、フェールオーバーが発生してサーバAのサービスaを引き継ぐ場合は、サーバBは固定ディスク3の代わりにサービスディスクDと論理的に接続されることが事前に分かっている。そこで、本実施形態では、フェールオーバー発生時の論理アドレスc1t0d0をあらかじめスクリプトファイル55として定義し、これをサーバBに格納するようにしている。

【0033】リンク処理部53は、サービスディスクDのオーナ権がサーバBに移行される前まで設定されていた論理アドレスを、スクリプトファイル55にリンク情報としてあらかじめ記述されている論理アドレスに書き換える処理を行う。すなわち、上記オーナ権移行処理部52によってサービスディスクDのオーナ権がサーバBに移行されたら、リンク処理部53は、リンクを張るための“ln”といったコマンドを発行して、上記リンク情報に基づいてc1t0d0という関連を認識し、実際にサーバBとサービスディスクDとのリンクをはる。

【0034】オーナ権リリース処理部54は、詳しくは後述するが、サービスディスクDのオーナ権をリリースする指令として、“autotrespass”といったコマンドを発行し、それまでサーバBが保持していたオーナ権の解放処理を行うものである。このオーナ権リリース処理部54は、例えば、故障したサーバAが復旧した際に、サービスディスクDのオーナ権をサーバAに返すために手動で起動される。

【0035】次に、上記実施形態のフェールオーバー管理システムの動作について説明する。図3は、サーバB側で行われる処理を簡単に示したフローチャートで、この処理手順は主にスクリプトファイル55の記述に従って実行される。サーバBは、デーモン監視部51及び監視ネットワーク4、5によってサーバAを監視し（ステップS101）、サーバAに障害が発生していないかどうかを判別する（ステップS102）。

【0036】サーバAに障害が発生していないと判別した場合、すなわち、二重化された監視ネットワーク4、5で通信が途絶えていないと認定した場合、ステップS101に戻って繰り返し監視を続ける。一方、サーバAに障害が発生していると判別した場合、すなわち、二重化された監視ネットワーク4、5の両方で通信が途絶えたことと認定した場合、オーナ権移行処理部52でオーナ権移行指令（コマンド“trespass”）を発行し、サービスディスクDのオーナ権をサーバAからサーバBに移行させる（ステップS103）。

【0037】続いて、リンク処理部53において、既に説明したように、コマンド“ln”を発行して、スクリプトファイル55に記述されているリンク情報に基づいて実際にサービスディスクDとのリンクをはる（ステップS104）。さらに、サービスディスクDをマウントして、サービスディスクDのファイルを実際に使用可能

な状態にする(ステップS105)。そして、それまでサービスaを提供していたサーバAと同じIPアドレスに基づいてネットワーク起動を行うとともに(ステップS106)、アプリケーション起動を行って(ステップS107)、サービスaの引き継ぎが終了する。

【0038】以上述べたようにして、サーバAに障害が発生した場合、そのサーバAが提供していたサービスaをすみやかにサーバBに引き継がせることができるので、サーバAのシステムダウンに伴う業務停止を免れることができる。

【0039】次に、上記のようにサーバBにサービスaが引き継がれた後、障害が取り除かれたサーバAに、どのようにしてサービスaを戻すかについて説明する。サーバAにサービスaを戻すには、サーバBにサービスaを引き継がせたのと同様に、サーバAでコマンド"trespass"を発行して、サービスディスクDのオーナー権を自己に移行させる方法が考えられる。しかし、サーバAでコマンド"trespass"を発行するためには、その時点でサーバAからはサービスディスクDが見えていないため、RAID1内のサーバA側にもそのコマンドを受け取るための固定ディスクが必要となる。そのため、サーバA、Bの両側で固定ディスクが必要となり、固定ディスクの数が多くなってしまう。特に、冗長性を持たせるためには、サーバA、Bそれぞれで2つ以上の固定ディスクを設けるのが望ましいことから、コストアップの要因となってしまう。

【0040】そこで、本実施形態では、図示しないキーボード操作等によって、サーバBのオーナー権リリース処理部54でオーナー権リリース指令(コマンド"autotrespass")を発行させるようにしている。このコマンド"autotrespass"が発行されると、サービスディスクDのオーナー権はサーバBからリリースされ、その後アクセスのあったサーバ側に順次自動的に移行することになる。したがって、サービスディスクDのオーナー権をサーバB側からリリースさせた後に、そのサービスディスクDに障害が取り除かれたサーバA側からアクセスするよう制御すれば、サーバAがサービスaを提供し、また、サーバBが待機中となる元の状態に復帰させることができる。そして、元の状態に復帰させたら、コマンド"autotrespass"の発行を止めればよい。

【0041】このように、一のサーバAを通常状態でサービスaを提供するプライマリサーバ、他のサーバBを待機系のサーバとすることにより、サーバAからサーバBの方向に対するフェールオーバーだけが発生するようにし、フェールオーバー完了後にサーバBからサーバAにサービスaを戻すときに、サーバB側からオーナー権リリース指令(コマンド"autotrespass")を発行するようすれば、サーバA側にコマンド"trespass"を受け取るための固定ディスクを設ける必要がなくなる。したがって、待機中のサーバB側に固定ディスク3を設けるだけ

で済み、固定ディスクの数が増えるのを避けて、コストダウンを図ることができる。

【0042】以上述べたように、本実施形態によれば、待機中のサーバBにおいて、フェールオーバーの発生時にサーバBに接続されるべきサービスディスクDの物理アドレスと論理アドレスとのひもづけをスクリプトファイル55によってあらかじめ定義しているため、サーバBがサービスaを引き継ぐ際に、そのスクリプトファイル55の記述に従って論理アドレスの設定を変更するだけ済み、SCSIのデバイスをスキャンしてその中からディスクを探してリンクをはるといった処理を行う必要がなくなる。したがって、全てのデバイスをスキャンしてOS上にデバイスファイルを再構築するという重い処理を行わなくて済むことから、サービスの引き継ぎ時間を大幅に短縮することができる。

【0043】なお、上記実施形態では本発明を非対称型のシステムに適用した例を説明したが、本発明を他のタイプのシステム、例えば対称型のシステムに適用してもかまわない。図4には、本発明を対称型のシステムに適用した場合の概略構成図を示す。対称型のシステムでは、サーバA、Bが、ネットワーク2にそれぞれ2回線で接続されている。

【0044】また、RAID1には、サービスaを提供するためのサービスディスクDaと、サービスbを提供するためのサービスディスクDbとが内蔵されている。これらサービスディスクDa、DbはいずれもサーバA、Bに物理的に接続されているが、通常は、サーバAからはサービスディスクDbが見えないように、また、サーバBからはサービスディスクDaが見えないように設定されている。したがって、サーバAは、1回線を使用してサービスディスクDaを用いたサービスaを提供するとともに、1回線でサービスbの待機中となっている。同じく、サーバBでも、1回線を使用してサービスディスクDbを用いたサービスbを提供するとともに、1回線でサービスaの待機中となっている。

【0045】サーバA、Bはいずれも、図2で示したブロック図の構成をとり、デモン監視部51と、オーナー権移行処理部52と、リンク処理部53と、オーナー権リリース処理部54と、スクリプトファイル55の記憶部とを有する。そして、サーバAに障害が発生した場合には、上記実施形態で説明したように、サーバBで、コマンド"trespass"を発行してサービスディスクDaのオーナー権を取得するとともに、スクリプトファイル55内のリンク情報に基づいてサービスディスクDaとのリンクをはって、そのサービスaを引き継ぐ。したがって、サーバBは、既存のサービスbとともに、引き継いだサービスaを2回線使用してネットワーク2に提供することになる。

【0046】同様に、サーバBに障害が発生した場合にも、サーバAでコマンド"trespass"を発行してサービ

スディスクD bのオーナー権を取得するとともに、スク립トファイル55内のリンク情報に基づいてサービスディスクD bとのリンクをはって、そのサービスbを引き継ぐ。したがって、サーバAは、既存のサービスaとともに、引き継いだサービスbを2回線使用してネットワーク2に提供することになる。すなわち、対称型のシステムでは、サーバA、Bが、状況によって本発明でいう一のサーバを構成したり、他のサーバを構成したりすることとなる。

【0047】なお、対称型システムの場合、通常はサーバA、BからそれぞれサービスディスクD a、D bが見える状態となっているので、一方のサーバがダウンしたときに他方のサーバからコマンド"trespass"を発行する対象となるディスクは必ず存在する。したがって、コマンド"trespass"を受け取るための固定ディスクを別に設ける必要はない。すなわち、サーバAに障害が発生して、サーバBでコマンド"trespass"を発行するときは、そのコマンドをサービスディスクD bで受け取ればよい。同様に、サーバBに障害が発生して、サーバAでコマンド"trespass"を発行するときは、そのコマンドをサービスディスクD aで受け取ればよい。

【0048】このような構成にした対称型のフェールオーバー管理システムでも、引き継いだサービスを元のサーバに戻すときに、サービスを引き継いだ側のサーバで、コマンド"autotrespass"を発行するようにしている。この場合も、コマンド"autotrespass"を発行するサーバからはサービスディスクD a、D bが見える状態となっているので、コマンド"autotrespass"を受け取るための固定ディスクを別に設けておく必要はない。

【0049】以上述べた対称型システムにおいては、サービスの引き継ぎ時間を大幅に短縮することができるという効果に加えて、サービスを引き継ぐ側のサーバが自己のサービスを行っている途中で全デバイスのスキャンを行うという重い処理をなくすことができ、サービスを引き継ぐ側のサーバにおける既存のサービスに支障が生じるのを防ぐこともできる。

【0050】ここまでは、2台のサーバを有するフェールオーバー管理システムについて説明したが、サーバの数は2台に限定されるものではない。例えば、図5に示すシステムでは、サーバAがサービスaを提供し、サーバBがサービスbを提供し、さらにサーバCがサービスa、bの待機中となっている。そして、サーバA、Bに障害が発生した場合には、既に説明したようにしてサーバCにサービスディスクD a、D bのオーナー権を移行させて、サーバCがそのサービスa、bを引き継ぐようにしている。

【0051】なお、上記実施形態では、サービスディスクDが本発明でいうデバイスを構成する。ただし、デバイスとしては、サービスディスクD以外の記憶装置でもよく、さらには記憶装置以外のものからなるデバイスで

あってもよい。

【0052】以上説明した各実施形態のフェールオーバー管理システムは、コンピュータのCPUあるいはMPU、RAM、ROM等で構成されるものあり、RAMやROMに記録されたプログラムが動作することによって実現できる。したがって、コンピュータが上記機能を果たすように動作させるプログラムを記録媒体に記録し、コンピュータに読み取らせることによって実現できるものである。記録媒体としては、CD-ROM、フロッピー（登録商標）ディスク、ハードディスク、磁気テープ、光磁気テープ、不揮発性のメモ리카ード等を用いることができる。

【0053】また、コンピュータが供給されたプログラムを実行することにより上述の実施形態の機能が実現されるだけでなく、そのプログラムがコンピュータにおいて稼働しているOS（オペレーティングシステム）あるいは他のアプリケーションソフト等と共同して上述の実施形態の機能が実現される場合や、供給されたプログラムの処理の全てあるいは一部がコンピュータの機能拡張ボードや機能拡張ユニットにより行われて上述の実施形態の機能が実現される場合にも、かかるプログラムコードは本発明の実施形態に含まれることは言うまでもない。

【0054】なお、上記いずれの実施形態も本発明を実施するにあたっての具体的な一例を示したに過ぎず、これによって本発明の技術的範囲が限定的に解釈されてはならない。すなわち、本発明は、その精神、またはその主要な特徴から逸脱することなく、様々な形で実施することができる。

【0055】

【発明の効果】以上述べたように本発明によれば、他のサーバが一のサーバのサービスを引き継ぐ際の物理アドレスと論理アドレスとのひもづけ情報を他のサーバにあらかじめ格納しているため、当該他のサーバでのサービスの引き継ぎ時に全てのデバイスをスキャンしてその中からディスクを探してリンクをはるといった重い処理を行う必要がなくなり、サービスの引き継ぎ時間を大幅に短縮することができる。また、他のサーバが他のサービスを提供しているような場合でも、スキャンという重い処理を行わなくて済むことから、既存の他のサービスに支障が生じるのを防ぐことができる。

【図面の簡単な説明】

【図1】本実施形態の非対称型のフェールオーバー管理システムの概略構成図である。

【図2】サーバBの機能構成を示すブロック図である。

【図3】サーバB側で行われる処理を示したフローチャートである。

【図4】本実施形態の対称型のフェールオーバー管理システムの概略構成図である。

【図5】3台のサーバを使用したフェールオーバー管理シ



システムを説明するための図である。

【図6】非対称型のフェールオーバー管理システムを説明するための図である。

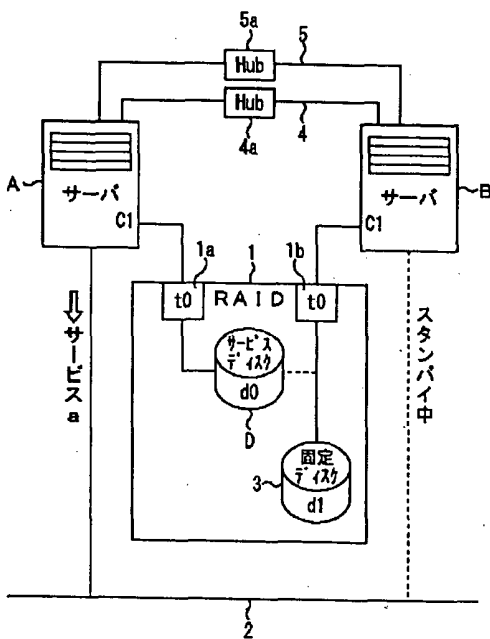
【図7】対称型のフェールオーバー管理システムを説明するための図である。

【符号の説明】

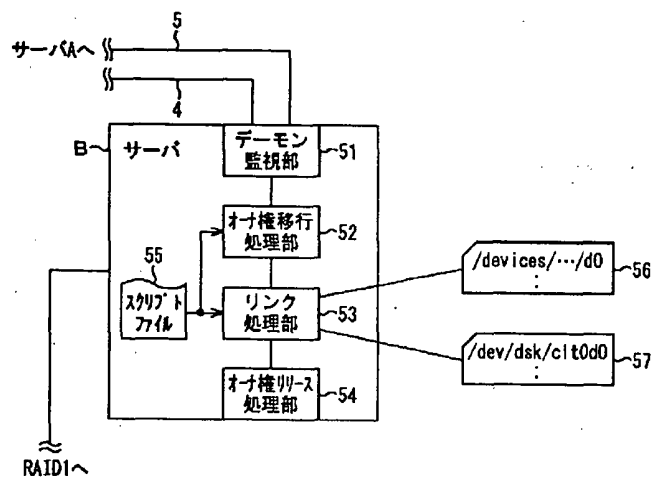
A、B、C サーバ  
D、Da、Db サービスディスク  
1 RAID  
1a、1b RAIDコントローラ

2 ネットワーク  
3 固定ディスク  
4、5 監視ネットワーク  
51 デーモン監視部  
52 オーナ権移行処理部  
53 リンク処理部  
54 オーナ権リリース処理部  
55 スクリプトファイル  
56 物理アドレス  
57 論理アドレス

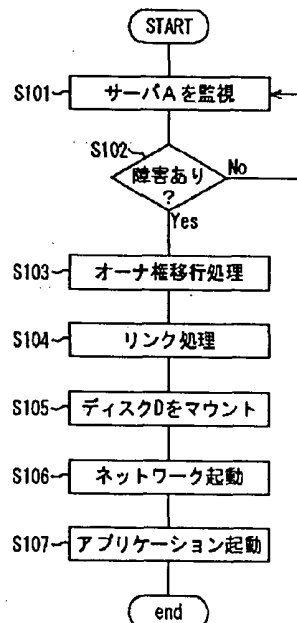
【図1】



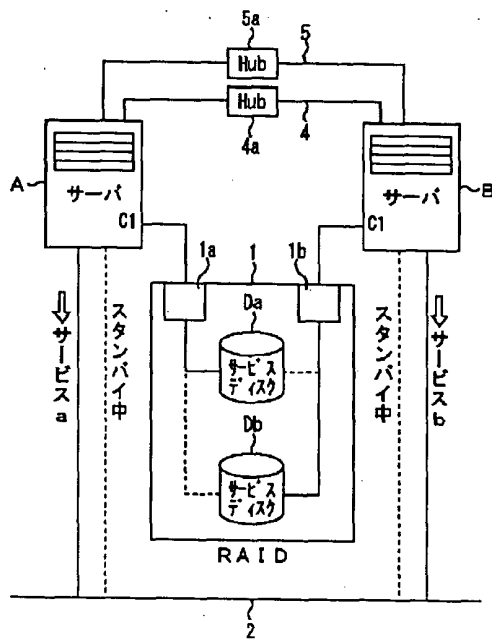
【図2】



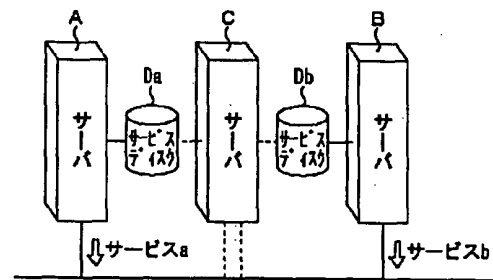
【図3】



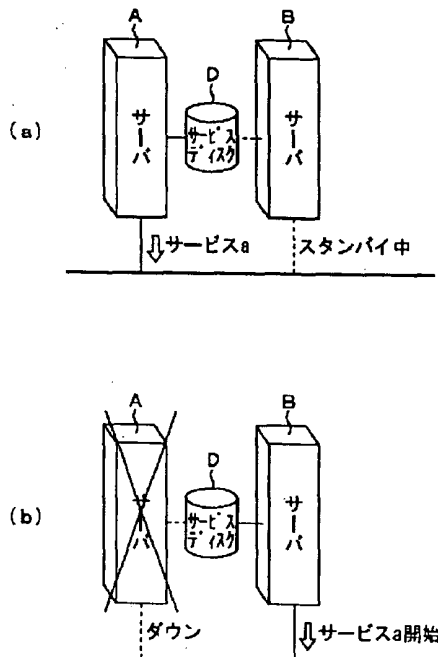
【図4】



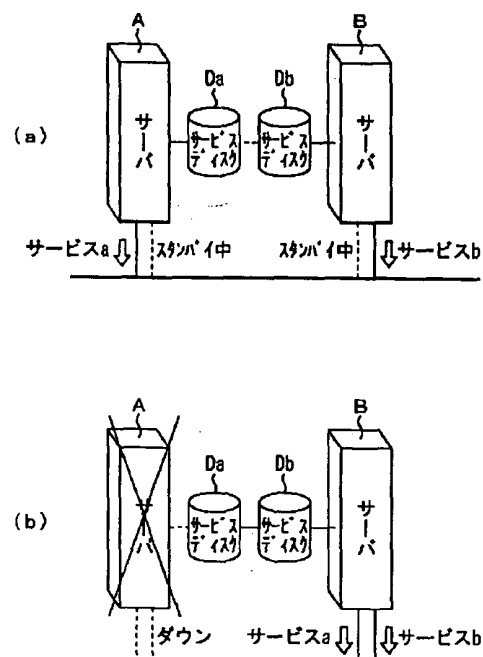
【図5】



【図6】



【図7】



フロントページの続き

Fターム(参考) 5B034 BB02 CC01 DD02 DD07

5B045 GG06 JJ02 JJ13 JJ26 JJ44

JJ46

5B083 AA08 BB03 CC01 CD11 CE01

DD01 DD09 EE02

5B089 GA11 GB02 JA33 KA12 KB06

KC15 KC21 MD01